

This is a repository copy of *Exploring implications of context specificity and cognitive load in residents*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/149771/>

Version: Published Version

---

**Article:**

Ratcliffe, Temple, McBee, Elexis, Schuwirth, Lambert et al. (5 more authors) (2017)  
Exploring implications of context specificity and cognitive load in residents. AMEE  
MedEdPublish.

<https://doi.org/10.15694/mep.2017.000048>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# Exploring Implications of Context Specificity and Cognitive Load in Residents

Temple A. Ratcliffe[2], Elexis McBee[3], Lambert Schuwirth[4], Katherine Picho[5], Cees P. M. van der Vleuten[6], Anthony R. Artino[5], Jeroen J. G. van Merriënboer[6], Jimmie Leppink[6], Steven J. Durning[5]

**Corresponding author:** Dr Temple A. Ratcliffe [ratcliffe@uthscsa.edu](mailto:ratcliffe@uthscsa.edu)

**Institution:** 2. UT Health San Antonio, 3. Naval Medical Center San Diego, 4. Flinders University, 5. Uniformed Services University of the Health Sciences, 6. Maastricht University

**Categories:** Medical Education (General), Research in Medical Education

Received: 10/03/2017

Published: 14/03/2017

## Abstract

**Introduction:** Context specificity (CS) refers to the variability in clinical reasoning across different presentations of the same diagnosis. Cognitive load (CL) refers to limitations in working memory that may impact clinicians' clinical reasoning. CL might be one of the factors that lead to CS. Although CL during clinical reasoning would be expected to be higher in internal medicine residents, CL's effect on CS in residents has not been studied.

**Methods:** Internal medicine residents watched a series of three cases portrayed on videos. Following each case, participants filled out a post-encounter form and completed a validated measure of CL.

**Results:** Fourteen residents completed all three cases. Across cases, self-reported CL was relatively high and there were small to moderate correlations between CL and performance in clinical reasoning ( $r$ 's = .43, -.33, -.23). In terms of changing CL across cases, the correlations between change in CL and change in total performance were statistically significant only in moving from case 1 to case 2 ( $r = -.54, p = .05$ ).

**Discussion and Conclusion:** Residents self-reported measurements of CL were relatively high across cases. However, higher CL was not consistently associated with poorer performance. We did observe the expected associations when looking at case-to-case change in CL. This relationship warrants further study.

**Keywords:** clinical reasoning

## Introduction

---

Clinical reasoning is a complex phenomenon (Durning, Artino, Schuwirth & van der Vleuten, 2013). Although research over the past few decades has identified frameworks that have improved our understanding of what constitutes clinical reasoning and how to assess it, significant questions remain such as what leads to context specificity (CS) (i.e., the finding that factors outside of a case's clinical content impact reasoning) (Durning et al 2012).

One framework applied to clinical reasoning is situated cognition. Situated cognition posits that clinical reasoning does not exist solely in the physician's mind, but rather is situated "in the specifics of the event" and is made up of interactions between a doctor, his or her patient and other encounter factors (Durning & Artino 2011). In this framework, context is essential, as changes in context may ultimately affect clinical reasoning performance. Through this lens, contextual factors are dimensions of an individual case that may pertain to the doctor, patient, encounter and/or their interactions. Contextual factors are proposed to affect reasoning within the situated cognition framework and through these interactions they could contribute to CS (contextual factors, in addition to the clinical content of a case, might lead to case to case performance variability).

Prior work on the effect of contextual factors on performance in both experts and residents (including a subset of these participants) (McBee et al 2015, McBee et al 2016), has demonstrated that introducing contextual factors decreased clinical reasoning performance in otherwise straightforward medical cases. The mechanism for this effect on performance is not well defined. One possible mechanism for the influence of contextual factors on performance is cognitive load (CL). Of note, previous work with think-aloud methodology provided evidence that limitations in CL (evidenced by decreased semantic competence on think alouds) may be an important factor in decreased performance in experts (Durning, Artino, Pangaro, van der Vleuten & Schuwirth, 2011).

In contrast with situated cognition, cognitive load theory (CLT) traditionally focuses on the level of the individual decision maker (i.e., clinician; although more recently external factors such as the physical environment are proposed to affect CL as well) (Choi, van Merriënboer & Paas, 2014). CLT explores the implications of known limitations of working memory. Due to these biologic limitations, doctors may only be able to manipulate three to four "chunks" of information in working memory at one time (van Merriënboer & Sweller, 2010). These limitations have implications for the clinical reasoning process. Contemporary CLT distinguishes between intrinsic and extraneous CL (Leppink, Paas, Van Gog, Van der Vleuten & Van Merriënboer, 2014). Intrinsic CL includes information essential to solving a clinical problem and reflects how complex a task is given the expertise of the participant. Extraneous CL results from present, but ultimately unnecessary, processes or information that is not essential to working through a clinical scenario.

CLT offers one explanation of how contextual factors may influence physician performance by affecting CL. Specifically, from a situated cognition perspective, we suspect contextual factors (if they are not relevant to the case) could mediate their effects on performance by increasing extraneous CL. Our prior work has suggested that the introduction of contextual factors to a clinical scenario could impact clinical performance by impacting extraneous CL (and thus overall CL) in participants (Durning et al, 2011, Durning, et al 2012.), however CL was not directly measured in these investigations. Instead, indirect evidence was inferred from decreased expert semantic competence during think alouds. In the current investigation, we wished to directly measure the association of CL and performance.

Because physicians typically see multiple patients a day, we believed that it was important to look at clinical reasoning performance and CL across a series of cases. So in addition to evaluating CL for a single case, we were

also interested in the change in CL across cases. Our specific research question was: In residents, are self-reported CL measurements associated with clinical reasoning as measured by a post-encounter form? We hypothesized that 1) CL would be high across cases in part due to the addition of contextual factors in the video cases even though the content of the cases was written at a straightforward level (as subsequently confirmed by a panel of experts). We also hypothesized that 2) CL would be negatively related to performance on the post-encounter form. Finally, we hypothesized that 3) sustained CL would result in declining performance.

## Methods

---

The present study was conducted between 2012-2013 with internal medicine residents at the Uniformed Services University and Brooke Army Medical Center. The study was approved by the IRB at the Uniformed Services University and then acknowledged and approved in memo by Brooke Army Medical Center IRB.

### Measures

**Cognitive Load.** CL was assessed by a single-item measure (Paas, Moreno & Brunken, 2010) that asks participants to rate their invested mental effort on the given task. This item is anchored on a 9-point Likert scale ranging from 1 (no effort at all) to 9 (very high effort).

**Clinical reasoning.** The study outcome was clinical reasoning performance and CL was used as a predictor variable. Clinical reasoning performance was assessed by a previously validated post encounter form (Durning, 2012A). This form assesses participants' approaches to obtaining information from patient histories and the physical exams; participants also provide leading and differential diagnoses, offer supporting data justifying the choice of a leading diagnosis, and list treatment plans for each case.

### Procedure

Participants viewed three video-recorded standardized patient encounters of three different straightforward disease presentations. Each video included a variable number of contextual factors that provided irrelevant information for purposes of diagnosis or therapy. The physician and patient in the cases were portrayed by actors and displayed one of three potential diagnoses: symptomatic type 2 diabetes mellitus, HIV, and colorectal cancer. Contextual factors were manipulated to facilitate inquiry into the impact of context on diagnostic and therapeutic reasoning performance. Contextual factors manipulated in these videos pertained to: 1) a patient with low English proficiency, 2) emotional volatility displayed as challenging the physician's credentials or 3) a combination of both of the aforementioned factors (see Table 1). After completing a post-encounter form for each case, participants completed the single-item measure of CL.

Table 1: Video contextual factors

Case	Contextual Factor	Diagnosis
1	Low English Proficiency	HIV
2	Emotional volatility	Colorectal cancer
3	Both	Diabetes mellitus

## Analysis

**Scoring diagnostic reasoning.** Participants' post-encounter forms were scored by two coders (SD, TR) using an established rubric. Each of the six sections (additional history, additional physical exam, problem list, differential diagnosis, leading diagnosis, and treatment plan) was scored such that correct, partially correct and incorrect entries were allocated (2, 1, and 0 points, respectively). We wanted to determine whether the results of the six could be pooled. Because the numbers of participants were low we performed four analyses of internal consistency using Cronbach's alpha. First we calculated alpha over all the items of the three cases ( $6 \times 3 = 18$  scores) which was .229; second we calculated alpha over the total scores of each section (6 scores) which was .180 and finally we calculated the alpha over the items within each case (6 items per case) which were .436; .130 and .171 for cases 1, 2 and 3 respectively. Therefore we concluded that in our data we were unable to detect any sign of specificity of the six sections higher than the expected case specificity and decided to sum the scores across all sub categories to yield a total value to denote clinical reasoning.

The relationship between diagnostic reasoning performance and CL was examined with Spearman's rank correlations using Stata12. However, in order to investigate whether a change in CL was associated with change in (1) total performance and (2) performance on presumed high cognitively demanding tasks, we created change scores for both CL, total performance, and performance on these presumed high CL tasks. This was done separately for all three cases. Then for each case, two  $\Delta$ CL were created by subtracting CL scores of one case from the other in a linear fashion (i.e., case 2- case 1 and case 3-case 2) for total performance and total performance on high CL task outcomes variables.

## Results

Fourteen participants completed three cases each. Participants' mean age was 30 years, 57% (8/14) were female, and the majority were in their second year of residency (21%, 3/14 PGY-1; 57%, 8/14 PGY-2; 21%, 3/14 PGY-3).

Self-reported measurements of CL averaged above 6 per case on a 9 point Likert scale (Table 2). Participants' performance varied across each section of each case as well as from one case to the next (Table 3). Overall across

cases, small to moderate, but inconsistent, correlations between CL and performance in clinical reasoning (scores aggregated within each case;  $r$ 's = .43, -.33, -.23). The correlation was positive only for case 1 and there was no significant change in CL in going from case 1 to case 3 (Table 2) ( $p > .05$ ). The expected negative correlations between CL and performance were seen only for case 2 and case 3.

In terms of  $\Delta$  CL across cases, the correlations between  $\Delta$  CL and change in total performance were statistically significant in moving from case 1 to case 2 ( $r = -.54, p = .05$ ), but small and not statistically significant in going from case 2 to case 3 ( $r = -.09, p > .05$ ). The correlation was relatively high for the  $\Delta$  CL going from case 1 to case 2 ( $r = -.63, p = .02$ ), but small and not statistically significant in going from case 2 to case 3 ( $r = -.28, p > .05$ ).

Table 2: Mean Cognitive Load Scores and correlations across cases

	Cognitive Load	
Case	Mean	SD
Case 1	6.5	1.27
Case 2	6.6	1.50
Case 3	7.0	1.56

Table 3: Performance Scores for clinical reasoning sub-categories, and related correlations with cognitive load

	Descriptive Statistics				Correlations						
	n	M	SD		1	2	3	4	5	6	7
Case 1											
Patient history	14	5.5	1.34		1	-.53*	-.37	-.61	-.7**	-.05	-.47
Patient exam	14	4.71	2.12			1	.24	.56	.29	-.07	.15
Differential dx	14	2.79	1.76				1	.63	.67**	-.12	.33
Lead dx	6	1.33	0.52					1	.38	1.0***	.77
Supporting data	13	2.69	2.98						1	-.04	.44
Treatment plan	7	1.43	0.53							1	.52
Cognitive load	14	6.43	1.09								1
Case 2											
Patient history	14	6.36	1.39		1	-.06	-.73**	--^	.27	.18	-.35
Patient exam	14	3.57	1.50			1	.15	--^	-.15	-.27	-.09
Differential dx	14	4.50	1.70				1	--^	.10	-.38	.12
Lead dx	7	2.00	0.00					--^	--	--	--
Supporting data	14	6.93	1.49						1	-.14	-.34
Treatment plan	11	1.36	0.67							1	.22
Cognitive load	14	6.57	1.40								1

Case 3											
Patient history	14	4.36	1.94		1	-.35	.22	-.51	-.49	-.15	.02
Patient exam	14	4.79	1.67			1	.13	.51	.55*	.23	-.02
Differential dx	14	4.07	1.54				1	-.16	.03	.36	-.29
Lead dx	9	1.44	0.53					1	.78*	.26	-.19
Supporting data	14	6.50	2.21						1	.39	-.31
Treatment plan	12	1.67	0.49							1	-.12
Cognitive load	14	7.07	1.38								1

Note. \*  $p < .05$  \*\*  $p < .01$  \*\*\*  $p < .001$

^= Range restriction and participants not listing an answer (missing variables) resulted in inability to calculate correlations.

## Discussion

As expected based on our prior work, and consistent with a situated cognition framework, participants' performance varied across cases despite cases being written as straightforward presentations of diseases. What situated cognition had not specifically addressed, however, is why this variation in performance may occur. Here we invoked CLT as it pertains to how individuals process information and perform. Through interactions with others and the environment, it is plausible that CL may mediate some of the case to case variation seen in prior work (Durning et al, 2011). CL, specifically extraneous CL, may at least partly moderate these interactions, and CL itself may be altered as a result of these.

As we hypothesized (hypothesis #1), self-reported measurements of CL were relatively high across cases (upper third of Likert scale). Since the clinical information presented was straightforward, this likely reflects the addition of contextual factors resulting in increased extraneous CL. However, our expected hypothesis (hypothesis #2) that higher overall CL would be associated with poorer performance was not consistently demonstrated. In fact, there was an unexpected positive correlation between CL and performance on the post-encounter form in case 1 (while trends in cases 2 and 3 were in the expected, negative direction). Certainly, different contextual factors may also have different effects, alone and in combination, on a given participant's extrinsic CL. This would be consistent with both the frameworks of situated cognition and CS in that performance is situation dependent and CL is influenced by patient, individual participant, and environmental factors. It also may simply reflect our small sample size and number of cases.

We did observe the expected associations (hypothesis #2) when looking at case-to-case  $\Delta$  CL. The potential reasons for this may also help to explain the varying associations between overall CL and performance noted above. The issue may be one of calibration. This finding suggests that while it may be difficult for participants to calibrate an initial CL response, a change in CL, in comparison with the prior case, may be a more accurate measurement of CL for each individual participant. Related to this point was the finding that while our participants' overall performance improved from case to case (despite sustained CL and contrary to hypothesis #3), each individual participant's performance became more closely related to his or her CL, as measured by case to case  $\Delta$  CL. The sequential nature of this change suggests that our participants were, on average, better able to calibrate load in relation to previous ratings, thus leading to the expected relationship between  $\Delta$  CL and performance.

We suspect that further study will clarify the relationship between CL and CS further. Taking a broader view, from



a situated cognition perspective, one could argue for the need to partially re-conceptualize CLT. While CL is based on the inherent limitations of working memory that exist within each participant's mind, situated cognition offers a framework to examine the impacts of CL beyond the individual. A similar reconceptualization has already been proposed in regards to CL and physical environmental factors (Choi et al, 2014). Similar to the effect of physical factors such as noise in a room or temperature level, the changes in CL that we hypothesize are at least in part due to the presence of contextual factors. In practice, the interactions between contextual factors and an individual's CL may not be limited to the individual. This dynamic interplay between doctor, patient and the environment, when viewed through the lens of CLT, offers an alternative viewpoint to examine context specificity and variations in performance. (A view that would only be partially captured by video-recorded, "static" cases, and would be expected to be even more pronounced in clinical practice where the relationships between contextual factors and physical environmental factors are dynamic).

Our study had several limitations, First, we used a single measure of overall CL. CL is a multi-dimensional construct. We expected that inserting contextual factors in our cases would increase CL, specifically extraneous CL. Future studies should use an instrument that can help us differentiate between intrinsic and extraneous cognitive load. A recent 10-item CL measurement scale (unfortunately published after our initial data collection was nearly complete) could allow exploration of the contributions of different types of CL in participants (Leppink, Paas, van der Vleuten, Van Gog, van Merriënboer, 2013). Additionally, we had a relatively small number of participants. Given that each participant had to commit a half-day to participate in our study, we believe that our sample size was reasonable. Third, performance with video cases may not replicate actual patient care scenarios. Fourth, our findings demonstrate associations and cannot prove causation. In the short term, the new 10-item self-report instrument that distinguishes between intrinsic and extraneous load could prove more effective in assessing CL compared to the one-item instrument used in this study. Eventually, more novel methods may provide additional insights to CL and its relationship to clinical reasoning performance.

## Conclusion

---

We believe that CLT will provide a reasonable lens through which to view the effects of contextual factors on clinical reasoning and suspect that further research will bear out this relationship more clearly. There are several potential implications from our work. First, interventions that reduce extraneous CL may improve clinical reasoning, especially for non-experts (e.g., resident physicians and medical students). Extraneous CL is likely reducible with practice. One strategy would be to educate resident physicians on CFs and their potential influence on performance (just like we explicitly instruct on different types of illness presentations for a disease). Alternatively, attention to CFs in the work environment that could be addressed at a systems level (e.g., EMR improvements) could also lead to performance gains. Finally, while not explicitly explored in this investigation, when viewed from a situated cognition framework as well as CLT, high levels of fatigue, sleepiness, and burnout would also likely have a detrimental influence on performance through impact on CL. Thus, future work could explore these potential associations.

## Take Home Messages

---

## Notes On Contributors

---

Temple Ratcliffe, MD is an Associate Professor/Clinical at University of Texas Health at San Antonio.



Elexis McBee, DO, MPH is an Assistant Professor at the Uniformed Services University and practices at the Naval Medical Center in San Diego.

Lambert Schuwirth, MD, PhD is a Professor at Flinders University School of Medicine.

Katherine Picho, PhD is an Assistant Professor at Uniformed Services University.

Cees van der Vleuten, PhD is a Professor at Maastricht University.

Anthony R. Artino, Jr., PhD is an Associate Professor at Uniformed Services University.

Jeroen J. G. van Merrienboer, PhD is a Professor at Maastricht University.

Jimmie Leppink, PhD is a Professor at Maastricht University.

Steven J. Durning, MD, PhD is a Professor at Uniformed Services University.

## **Acknowledgements**

---

The authors would like to acknowledge Amy Flanagan Risdal for her contributions to the data collection for this study.

## **Bibliography/References**

---

Durning SJ, Artino AR. 2011. Situativity theory: a perspective on how participants and the environment can interact: AMEE guide no. 52. *Med Teach*. 33:188-199.

<https://doi.org/10.3109/0142159X.2011.550965>

Durning SJ, Artino AR, Boulet JR, Dorrance K, van der Vleuten C, Schuwirth L. 2012. The impact of selected contextual factors on experts' clinical reasoning performance (does context impact clinical reasoning performance in experts?) *Adv Health Sci Educ Theory Pract*. 17:65-79.

<https://doi.org/10.1007/s10459-011-9294-3>

Durning SJ, Artino AR, Boulet J, La Rochelle J, van der Vleuten C, Arze B, Schuwirth L. 2012A. The feasibility, reliability, and validity of a post-encounter form for evaluating clinical reasoning. *Med Teach*. 34:30-37.

<https://doi.org/10.3109/0142159X.2011.590557>

Durning S, Artino AR Jr, Pangaro L, van der Vleuten CP, Schuwirth L. 2011. Context and clinical reasoning: understanding the perspective of the expert's voice. *Med Educ*. 45:927-38.

<https://doi.org/10.1111/j.1365-2923.2011.04053.x>

Durning SJ, Artino AR Jr, Schuwirth L, van der Vleuten C. 2013. Clarifying assumptions to enhance our understanding and assessment of clinical reasoning. *Acad Med* 88:442-8.

<https://doi.org/10.1097/ACM.0b013e3182851b5b>

Choi HH, van Merriënboer JJG, Paas F. 2014. Effects of the physical environment on cognitive load and learning: towards a new model of cognitive load. *Educational Psychology Review*. 26:225-244.

<https://doi.org/10.1007/s10648-014-9262-6>

Leppink J, Paas F, van der Vleuten C, Van Gog T, van Merriënboer JJG. 2013. Development of an instrument for measuring different types of cognitive load. *Behav Res*. 45:1058-1072.

<https://doi.org/10.3758/s13428-013-0334-1>

Leppink J, Paas F, Van Gog T, Van der Vleuten CPM, Van Merriënboer JJG. 2014. Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*. 30:32-43.

<https://doi.org/10.1016/j.learninstruc.2013.12.001>

McBee E, Ratcliffe T, Goldszmidt M, Schuwirth L, Picho K, Artino AR Jr, Masel J, Durning SJ. 2016. Clinical Reasoning Tasks and Resident Physicians: What Do They Reason About? *Acad Med*. Jul;91(7):1022-8.

<https://doi.org/10.1097/acm.0000000000001024>

McBee E, Ratcliffe T, Picho K, Artino AR Jr, Schuwirth L, Kelly W, Masel J, van der Vleuten C, Durning SJ. 2015. Consequences of contextual factors on clinical reasoning in resident physicians. *Adv Health Sci Educ Theory Pract*. Dec;20(5):1225-36.

Paas JL, Moreno R, Brunken R. 2010. *Cognitive load theory*. New York (NY): Cambridge University Press.

van Merriënboer JJG, Sweller J. 2010. *Cognitive load theory in health professional education: design principles and strategies*. *Med Educ*. 44:85-93.

<https://doi.org/10.1111/j.1365-2923.2009.03498.x>

## Appendices

---

## Declarations

---

*The author has declared that there are no conflicts of interest.*

*This has been published under Creative Commons "CC BY 4.0" (<https://creativecommons.org/licenses/by-sa/4.0/>)*

---

AMEE MedEdPublish: rapid, post-publication, peer-reviewed papers on healthcare professions' education. For more information please visit [www.mededpublish.org](http://www.mededpublish.org) or contact [mededpublish@dundee.ac.uk](mailto:mededpublish@dundee.ac.uk).